DOCUMENT RESUME

ED 114 420 TM 004 914

AUTHOR Priest, Robert F.; Butler, Richard P.
TITLE Don't Use a Contingency Coefficient, Use Gamma.
PUB DATE [Apr 75]
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS Comparative Analysis; *Computer Programs; *Correlation; *Hypothesis Testing; Prediction; *Sampling; *Tests of Significance
IDENTIFIERS Chi Square; Contingency Coefficient; Gamma Coefficient

ABSTRACT

The advantages of Goodman's Gamma, a measure of association, are discussed in reference to the Pearson coefficient of contingency. Both theoretical and practical advantages and disadvantages are discussed. An empirical comparison of the two measures shows that gamma detects significant relationships which chi square does not, and that gamma is applicable to cases where chi square is not. (Author)

DON'T USE A CONTINGENCY COEFFICIENT, USE GAMMA

A paper presented at the 1975 AERA
Annual Meeting, Washington, DC

by

Robert F. Priest and Richard P. Butler
Office of the Director of Institutional Research
United States Military Academy
West Point, New York 10996

Abstract

The advantages of Goodman's Gamma, a measure of association, are discussed in reference to the Pearson coefficient of contingency. Both theoretical and practical advantages and disadvantages are discussed. An empirical comparison of the two measures shows that gamma detects significant relationships which chi square does not, and that gamma is applicable to cases where chi square is not.

DON'T USE A CONTINGENCY COEFFICIENT, USE GAMMA

In spite of the many warnings and difficulties (e.g., Lewis and Burke, 1949) involved in using chi square and its related measure of association, the Pearson contingency coefficient, these statistics are still used to test hypotheses about order. In several widely used statistics textbooks, the examples used to illustrate a contingency coefficient were two ordinal variables rather than two nominal variables. We think this practice is misleading. Perhaps usage of the contingency coefficient stems from Walker and Lev's (1953, p. 287) outdated statement that: "This is not a very satisfactory measure of relationship, but under the circumstances no better measure is available." However, the situation has changed since they wrote, and it is the thesis of this paper that researchers should no longer use the chi square and contingency coefficient for tests on ordinal variables.

Goodman and Kruskal's (1954) gamma coefficient validly avoids the difficulties that come with using the chi square and contingency coefficient on ordinal data. The gamma is a measure of association for determining the relationship between two ordinally scaled variables. Sophisticated computer programs, such as the Statistical Package for the Social Sciences (SPSS) can compute this measure for contingency tables (Nie, Bent, & Hull, 1970).

Nevertheless, users of the SPSS or other packages may be uneasy about using the gamma coefficient to measure ordinal relationships because they lack a convenient significance test. Although Goodman and Kruskal (1963) have worked out the sampling theory, the computation is complex and has <u>not</u> been incorporated into standard computer program packages such as the SPSS. However, we have developed a computer program to compute gamma, its standard error, its 95% confidence interval, and therefore its significance or nonsignificance. We will be happy to send a documented 11 page user's manual, complete with FORTRAN listing, to persons who write us.

Now that the <u>practical</u> difficulty has been solved, are there any <u>theoretical</u> reasons to prefer the gamma coefficient over the contingency coefficient? The advantages seem to lie with the gamma, because it has a simple and clear interpretation as a measure of monotonic association between two ordered variables. Gamma simply reflects the percentage of sample pairs which are similarly ordered on two variables. More precisely, if two persons are selected at random they are either tied on one of the two variables (x and y), or one person has a higher score on both variables, or one person is higher on x, with the other being higher on y. Discounting tied cases, the gamma coefficient measures the difference between the percentage of similarly ordered pairs. Or, it tells us the proportionate excess of concordant over discordant pairs among all pairs which are fully discriminated or fully ranked. Furthermore, gamma has directionality varying from -1 to +1, and you don't have to square a gamma to interpret it.

In contrast, the contingency coefficient does not have a simple interpretation. Its square does not measure explained variance, because in comparing two ordinally scaled variables the concept of variance is not meaningful. It has no sign; its upper limit varies with the number of rows and columns in the contingency tables; and it is not directly comparable to r, rho, tau, or any other correlation coefficient. In fact, it is possible for one to find a large chi square and contingency coefficient in situtations where the product-moment correlation is zero because of non-monotonic, non-linear relationships. This is not what most researchers usually want when they look for a measure of association. Of course, if that is what you want, by all means, use a contingency coefficient. Because of this insensitivity to ordering, the contingency coefficient used as an inferential tool is less powerful against population hypotheses of monotonic correlation than would be a test designed with such hypotheses specifically in mind. Therefore, when some monotonic correlation exists in the population, the traditional contingency coefficient test is less likely to reject the null hypothesis than, for example, a test for the significance of gamma, which is designed specifically for such alternative hypotheses.

We have recently compared the value of gamma and the conteingency coefficient as methods of describing the association between measures of leadership at the U.S. Military Academy and four criteria of post-Academy officer performance (Butler, 1973). The study was based on the Classes of 1961 through 1965. In all, 112 contingency tables, either 3x5 or 2x5, were developed. For 34 of these tables, it was impossible to compute a chi square and the contingency coefficients because the expected frequencies were too small. In 18 others, it was necessary to combine columns and rows. In contrast to this, the cell N's were too small to evaluate the significance of gamma in only 6 of the tables. This is so because of the sampling theory developed by Goodman and Kruskal (1963), which allows a researcher to evaluate the significance of a gamma for a wide range of sample sizes. Goodman and Kruskal developed large sample theory for the gamma coefficient, and its standard errors and significance tests are based upon the findings that gamma has an approximately normal sampling distribution for a sufficiently large sample size. From sampling experiments reported by them and by Rosenthal (1966), it is apparent that large sample statistics can be safely applied when there are an average of four cases for every cell in the table. For example, in a 3x4 contingency table, a sample size of 50 is adequate. Goodman and Kruskal (1963) also give a "conservative" standard error which can be used when the sample size is smaller than an average of four cases per cell. When the average frequency per cell is between two and four, the "conservative" standard error apparently produces acceptable results. Our computer program computes the appropriate standard error, depending on sample size. In our study, when both statistics were computable, nearly 70% of the comparisons showed that gamma was larger than the contingency coefficient, often by a fairly sizable amount. There were seven cases where the gamma coefficient was significant at the 5% level, whereas the contingency coefficient was not significant. As discussed earlier, the probable reason for this is that gamma is more sensitive to order. Furthermore, the gamma allowed a clearer interpretation, indicating the size of the monotonic association between two ordered variables. However, a significant contingency coefficient showed that there was a significant association among the variables, but not whether the relationship was monotonic, curvilinear, hyperbolical, etc. Further tests would have to be made to determine the type of relationship when using the contingency coefficient.

Two other recent studies have compared the sampling stability of the contingency coefficient with other measures. Whitney (1972) compared the chi square statistic and a rank order coefficient developed by Kendall. He showed that the rank order test is more powerful, provided that the underlying relationship between the two variables is linear or monotonic. Sarndal (1974) evaluated fifteen measures of association and concluded that the coefficient of contingency was among the most biased (pp. 185-186). In conjunction with data from the present research, one can conclude that the contingency coefficient is <u>never</u> the coefficient of choice.

REFERENCES

Butler, R. P. Aptitude for the service rating as a predictor of junior officer performance. U.S. Military Academy, Office of the Director of Institutional Research, West Point, New York, June 1973.

Goodman, L. A., & Kruskal, W. H. Measures of association for cross-classifications. Journal of the American Statistical Association, 1954, 49, 732-764.

Goodman, L. A., & Kruskal, W. H. Measures of association for cross-classifications: III, approximate sampling theory. Journal of the American Statistical Association, 1963, 58, 310-364.

Lewis, D., and Burke, C. J. The use and the misuse of the chi square test. Psychological Bulletin, 1949, 46, 433-489.

Nie, N., Bent, D., & Hull, C. H. SPSS: Statistical Package for the Social Sciences. New York: McGraw-Hill, 1970.

Rosenthal, I. The distribution of the sample version of the measure of association, Gamma. Journal of the American Statistical Association, 1966, 61, 440-454.

Sarndal, C. E. A comparative study of association measures. Psychometrika, 1974, 39, 165-187.

Walker, H. M., & Lev, J. Statistical Inference. New York: Holt, 1953.

Whitney, D. S and χ^2 tests of association: An empirical comparison. American Educational Research Journal, 1972, 9, 113-122.

DEPARTMENT OF THE ARMY
UNITED STATES MILITARY ACADEMY
WEST POINT, NEW YORK 10996

MAOR

MEMORANDUM FOR: RESEARCHERS AND PROGRAMMERS
OFFICE OF THE DIRECTOR OF INSTITUTIONAL RESEARCH

SUBJECT: Program GAMMA

1. Purpose: To compute GAMMA, a measure of association.

2. Background:

a. Goodman and Kruskal (1963) proposed GAMMA, measure of association, which reflects the degree of order in a contingency table. Unlike the product moment correlation coefficient, which is based on the principle of least squares the GAMMA coefficient optimizes the prediction of orde . The logic of the GAMMA coefficient is as follows: if two persons are selected at random, they are either tied on one of the two variables (x and y), or one person has a higher score on both variables, or one person is higher on x with the other higher on y. Discounting tied cases, the two variables are positively associated if there is a high probability for variables x and y to rank order two persons in similar ways. The two variables have a negative relation if variable x tends to rank-order persons opposite to the way variable y rank-orders them. The GAMMA coefficient is based on a direct count of the number of similarly rank-ordered pairs (S) minus the number of dissimilarly rank-ordered pairs (D).

$$\text{GAMMA} = \frac{S - D}{S + D}$$

b. The GAMMA coefficient has a rationale which is similar to the rationale for the tau coefficient, proposed by Kendall (1970). The tau coefficient is influenced by tied scores, whereas GAMMA is not. GAMMA is a particularly valuable measure to use, therefore, when there are many tied scores, such as one finds, in contingency tables.

c. Goodman and Kruskal (1963) developed large sample theory for the GAMMA coefficient. Standard errors and significance tests are based upon the fact that GAMMA has an approximately normal sampling distribution for a sufficiently large sample size. From sampling experiments reported by them and by Rosenthal (1966), it is apparent that large sample statistics can be safely applied when there are an average of 4 cases for every cell in the table. For example, in a 3x4 contingency table, a sample size of 50 is adequate. For a 5x5 table, N should be at least 100.

MAOR Memorandum For:
SUBJECT: Program GAMMA

d. Goodman and Kruskal (1963) also give a "conservative" standard error which can be used when the sample size is small. When the average frequency per cell is between 2 and 4, the "conservative" standard error apparently produces acceptable results.

3. <u>Description of the Program</u>:

Inputs: The program reads contingency tables (up to 20x20).

Outputs: and (1) prints each with a label
(2) computes GAMMA
(3) if the sample size is large enough computes the standard error and
(4) a 95% confidence interval for the population GAMMA.

The 95% confidence interval simultaneously tests a number of hypotheses about GAMMA. If the confidence interval does not include 0.00, then you can reject the usual null hypothesis that GAMMA = 0.00. In other words, when the confidence interval includes only positive non-zero values, you have a "significant" positive association. At the same time, if the confidence interval goes from .7 to .2, you may reject the hypotheses that the two variables correlate more than .7.

4. <u>How to Use the Program</u>: First, sort all the tables on the same size (same number of rows and columns) into one contiguous group.

Table Size Card:

Col 1-3 NR - Punch the number of rows in the contingency table right justify.*

Col 4-6 NC - Punch the number of columns in the contingency table, right justify.*

Col 7-9 NT - If there are a series of tables with the same number of rows and columns, punch the number of such tables. If you only have one RxC table, leave Col 7-9 blank.

Contingency Tables: Each table will require 2 or more cards. Each card has the same basic format.

*Right Justify: If you have a 3 digit number, leave two blank spaces, then punch the 3 digit into the right-most part of the 5 digit field. In general, if you have a short number to go into a large space, move the number over to the right side of the field, leaving the left side of the field blank.

The table should be arranged so that (reading left to right) each successive column represents successively higher values of variable x, and each successive row (moving down from the top row of the table) represents successively higher values of variable y. Thus, the first card you punch begins with the lowest ranked cell in the entire table; the last card you punch ends with the highest ranked cell. In the data deck, place the cards in the order punched, with top row (lowest y-values) first. For example, in the following table:

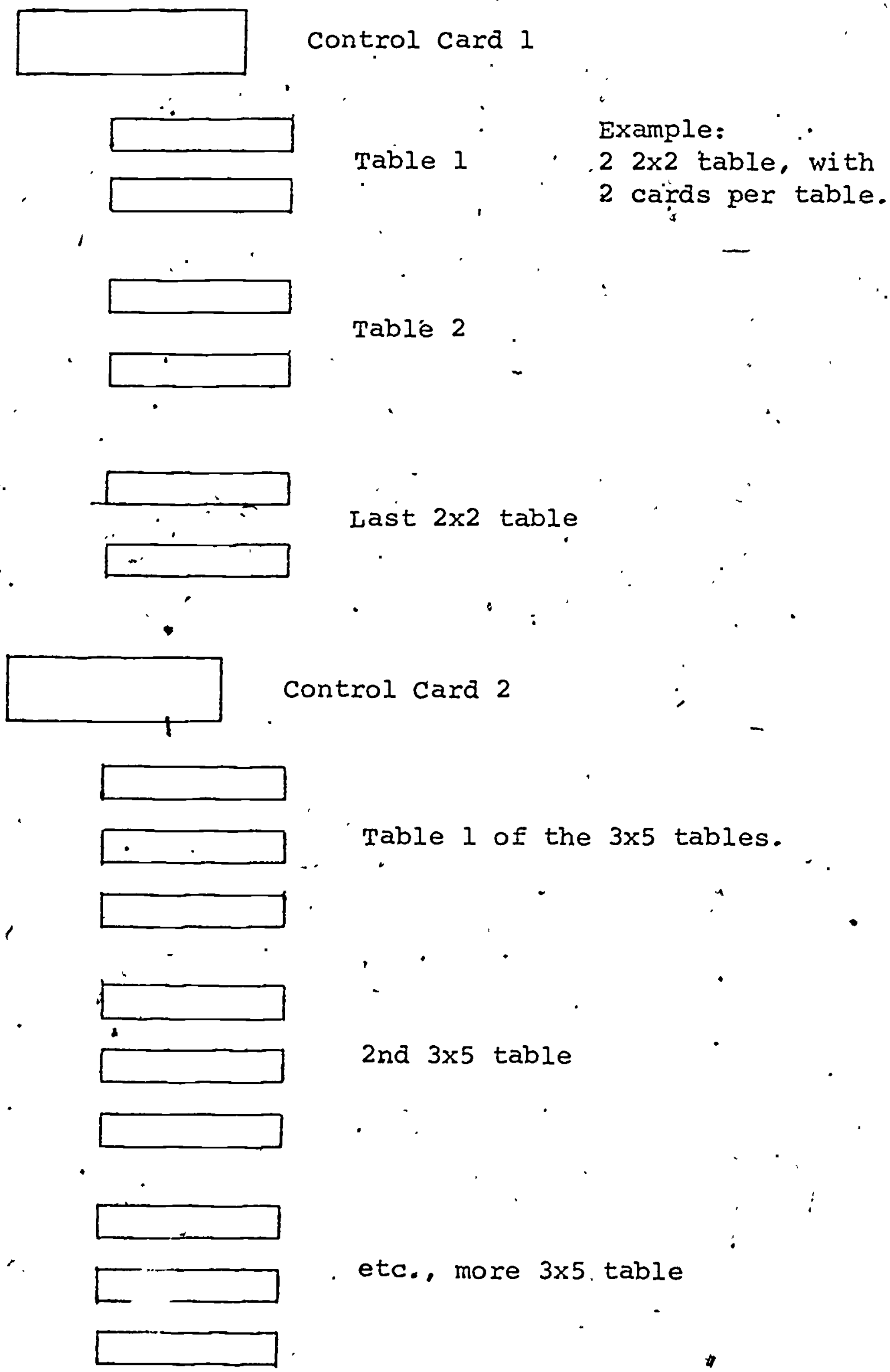
		X			
		Low	Med	Hi	
	Low	10	5	0	1st Card
Y	Med	5	11	4	2nd Card
	Hi	0	3	12	3rd Card

There are 10 scores with low scores on both X and Y; 12 scores with high scores on both X and Y; and so on. Start each new card with a label, and punch the call frequencies within one row of the contingency table, allowing five digits for each call frequency. Start each new row with a new card.

FORMAT FOR EACH ROW OF EACH TABLE

Col 1-5	Punch a 5 character label into each card. This label identifies the table for you in the output.
COL 6-10	Punch Row i, Col 1 frequency, right justify
COL 11-15	Punch Row i, Col 2 frequency, right justify
COL 16-80	Punch Row 1, Col 15, frequency, right justify

Symbolic Deck Order:



EXAMPLE OF JOB SET-UP

(A) Control Cards Which Preceed Program

```
0001 $ SNUMB    XZ971,40
0002 $ IDENT    ,MAOR/DL
0003 $ LIMITS   02,25K,,1000
0004 $ OPTION   FORTRAN,MAP
0005 $ FORTY
0006$ INCODE    USMA
```

GAMMA Program

```
C      GOODMAN-KRUSKALS GAMMA
C   CODED BY PRIEST-MAOR OCTOBER 1973
C  THIS PROGRAM PRODUCES A MEASURE OF ORDER ASSOCIATIONCALLED GAMMA
C  THE GAMMA COEFFICIENT WAS   PROPOSED BY GOODMAN AND KRUSKAL
C  IT MEASURES THE CONDITIONAL PROBABILITY THAT TWO RANDOMLY CHOSENPERSO  NS
C  WILL RANK IN THE SAME WAY ON VARIABLES X AND Y, MINUS THE PROBABILITY
C  THAT THEY WILL RANK IN OPPOSITE WAYS ON VARIABL X AND VARIABL Y
C  GIVEN THAT THEY ARE NOT TIED ON EITHER VARIABLE.
C     THE PROGRAM COMPUTES THE GAMMA COEFFICIENT
C                    ITS STANDARD ERROR
C                 AND A .95 PERCENT CONFIDENCE INTERVALAL FOR THE
C                         POPULATION GAMMMA
C          IF THE SAMPLE IS LARGE ENOUGH      THE BEST ASYMPTOTIC
C              STANDARDAR ERROR WILL BE O  COMPUTED
CC                   IF THE SAMPLE SIZE IS SMALL, THEN
C                THE PROGRAM COMPUTES A"CONSERVATIVE" STANDARD ERROR,
C                     IT IS THE UPPER BOUND FOR THE MORE PRECISE STANDARD  ERROR
C                    FORMULA.
C  IF THE SAMPLE SIZE IS TOO SMALL, THE PROGRAM PRINTS A MESSAGETO THAT  EFFE
C          -CT
C   HOW TO USE THIS PROGRAM
C   FIRST OF ALL, THERE IS ONE CONTROL CARD FOR EACH TABLE
C   OR SERIES OF TABLES.  THE CONTROL CARD TELLS THE PROGRAM THE NUMBER  OF
C    ROWS, THE NUMBER OF COLLUMNS IN THE NEXT TABLE OR SERIES OF TABLES
C   THE FIRST CONTROL CARD ALSO TELLS THE PROGRAMTHE NUMBER OF TABLES HA VE TH
C    SAME DIMENSION.  THESES THREE PIECES OF INFORMATION(PARAMETERS) ARE  CALL
C     NR, NC, NT BY THE PROGRAM.  PUNCH THEM INTO A SINGLE CARD
C   AND ALLOW 3 DIGITS FOR EACH NUMBER.  RIGHT JUSTIFY EACH NUMBER PUNCHED,
C
C    SECOND
C    ALWAYS PUNCH A FIVE DIGIT IDENTIFICATION NUMBER AT THE BEGINNING OF  EACH
C   CARD.  THIS FIVE DIGIT IDENTIFICATION NUMBER WILL BE USED TO HELP YOU
C   IDENTIFY THE PARTICULAR TABLE WHICH PRODUCED THE GAMMA
C   THEN CONTINUE PUNCHING EACH CARD WITHTHE CELL FREQUENCIES IN ONE ROW  OF
C   THE TABLE.  ALLOW FIVE DIGITS FOR EACH CELL FREQUENCY, AND RIGHT JUSTIFY
C  EACH NUMBER IF THERE ARE FEWER DIGITS IN THE CELL FREQUENCY,.
C    START EACH NEW ROW OF THE TABLE WITH A NUEWCARD OR SET OF CARDS
C   EACH TABLE SHOULD BE PUNCHED AS A SERIES OF 5 DIGIT NUMBERS
C
C
      DIMENSION X(20,20)
      DIMENSION S(20,20),D(20,20),IREG(4,4),Y(2)
 12   CONTINUE
      READ(11,100,END=99) NR,NC,NT
      IF(NT.EQ.0)NT=1
      DO 401 NTABLS=1,NT
C
C     NR = NO OF ROWS; NC = NO COLUMNS
C
      DO 400 I=1,NR
      READ(11,101,END=99)AD,(X(I,J),J=1,NC
    1 FORMAT(A5,10F5.0)
```

```
GOODMAN-KRUSKALS GAMMA.
CODED BY PRIEST-MAOR OCTOBER 1973
THIS PROGRAM PRODUCES A MEASURE OF ORDER ASSOCIATIONCALLED GAMMA
THE GAMMA COEFFICIENT WAS   PROPOSED BY GOODMAN AND KRUSKAL
IT MEASURES THE CONDITIONAL PROBABILITY THAT TWO RANDOMLY CHOSENPERSO NS
WILL RANK IN THE SAME WAY ON VARIABLES X AND Y, MINUS THE PROBABILITY
THAT THEY WILL RANK IN OPPOSITE WAYS ON VARIABL X AND VARIABL Y
GIVEN THAT THEY ARE NOT TIED ON EITHER VARIABLE.
 THE PROGRAM COMPUTES THE GAMMA COEFFICIENT
              ITS STANDARD ERROR
          AND A  95 PERCENT CONFIDENCE INTERVALAL FOR THE
                POPULATION GAMMMA
       IF THE SAMPLE IS LARGE ENOUGH      THE BEST ASYMPTOTIC
         STANDARDAR ERROR WILL BE O  COMPUTED
C             IF THE SAMPLE SIZE IS SMALL, THEN
          THE PROGRAM COMPUTES A"CONSERVATIVE" STANDARD ERROR;
               IT IS THE UPPER BOUND FOR THE MORE PRECISE STANDARD  ERROR
              FORMULA
IF THE SAMPLE SIZE IS TOO SMALL, THE PROGRAM PRINTS A MESSAGETO THAT EFFE
   -CT
HOW TO USE THIS PROGRAM.
FIRST OF ALL, THERE IS ONE CONTROL CARD FOR EACH TABLE
OR SERIES OF TABLES.  THE CONTROL CARD TELLS THE PROGRAM THE NUMBER  OF
 ROWS, THE NUMBER OF COLLUMNS IN THE NEXT TABLE OR SERIES OF TABLES
THE FIRST CONTROL CARD ALSO TELLS THE PROGRAMTHE NUMBER OF TABLES HA VE TH
 SAME DIMENSION.  THESES THREE PIECES OF INFORMATION(PARAMETERS) ARE  CALL
  NR, NC, NT BY THE PROGRAM.  PUNCH THEM INTO A SINGLE CARD
 AND ALLOW 3 DIGITS FOR EACH NUMBER.  RIGHT JUSTIFY EACH NUMBER PUNCHED.

 SECOND
 ALWAYS PUNCH A FIVE DIGIT IDENTIFICATION NUMBER AT THE BEGINNING OF  EACH
CARD.  THIS FIVE DIGIT IDENTIFICATION NUMBER WILL BE USED TO HELP YOU
IDENTIFY THE PARTICULAR TABLE WHICH PRODUCED THE GAMMA
THEN CONTINUE PUNCHING EACH CARD WITHTHE CELL FREQUENCIES IN ONE ROW  OF
THE TABLE.  ALLOW FIVE DIGITS FOR EACH CELL FREQUENCY, AND RIGHT JUSTIFY
EACH NUMBER IF THERE ARE FEWER DIGITS IN THE CELL FREQUENCY.
START EACH NEW ROW OF THE TABLE WITH A NUEWCARD OR SET OF CARDS
EACH TABLE SHOULD BE PUNCHED AS A SERIES OF 5 DIGIT NUMBERS

      DIMENSION X(20,20)
      DIMENSION S(20,20),D(20,20),IREG(4,4),Y(4)
   12 CONTINUE
      READ(11,100,END=99) NR,NC,NT
      IF(NT.EQ.0)NT=1
      DO 401 NTABLS=1,NT

      NR = NO OF ROWS,  NC = NO COLUMNS

      DO 400 I=1,NR
      READ(11,101,END=99)AD,(X(I,J),J=1,NC)
  101 FORMAT(A5,10F5.0)
```

PROGRAM LIST (1)

```
      PRINT 107,AD,(X(I,J),J=1,NC)
  107 FORMAT(1H ,A6,10F8.1)
  400 CONTINUE
      TX=0
      DO 33 I=1,NR
      DO 33 J=1,NC
   33 TX=TX+X(I,J)
      RC=NR*NC
C     IREG IS A DEVICE FOR CONTROLLING OPERATIONS ABOUT A PIVOT IN A MATRI)
      IREG(1,1) = 1
      IREG(1,3) = 1
      IREG(2,1) = 1
      IREG(3,3) = 1
      IREG(2,4) = NC
      IREG(3,2) = NR
      IREG(4,2) = NR
      IREG(4,4) = NC
  100 FORMAT(3I3)
C
C     START
C
C
C
       DO 1 I=1,NR
      IREG(1,2) = I-1
      IREG(2,2) = I-1
      IREG(3,1) = I+1
      IREG(4,1) = I+1
       DO 2 J=1,NC
      IREG(1,4) = J-1
      IREG(2,3) = J+1
      IREG(3,4) = J-1
      IREG(4,3) = J+1
        DO 4 K=1,4
       Y(K)=0.
      IA = IREG(K,1)
      IB = IREG(K,2)
      JA = IREG(K,3)
      JB = IREG(K,4)
      IF(IA.GT.IB) GO TO 4
      IF(JA.GT.JB) GO TO 4
         DO 5 II=IA,IB
      DO 5 JJ=JA,JB
    5 Y(K)=Y(K) + X(II,JJ)
    4 CONTINUE
      S(I,J)=Y(1)+Y(4)
      D(I,J)=Y(2)+Y(3)
    2 CONTINUE
    1  CONTINUE
      PS=0
      PD=0
      PSS=0.
```

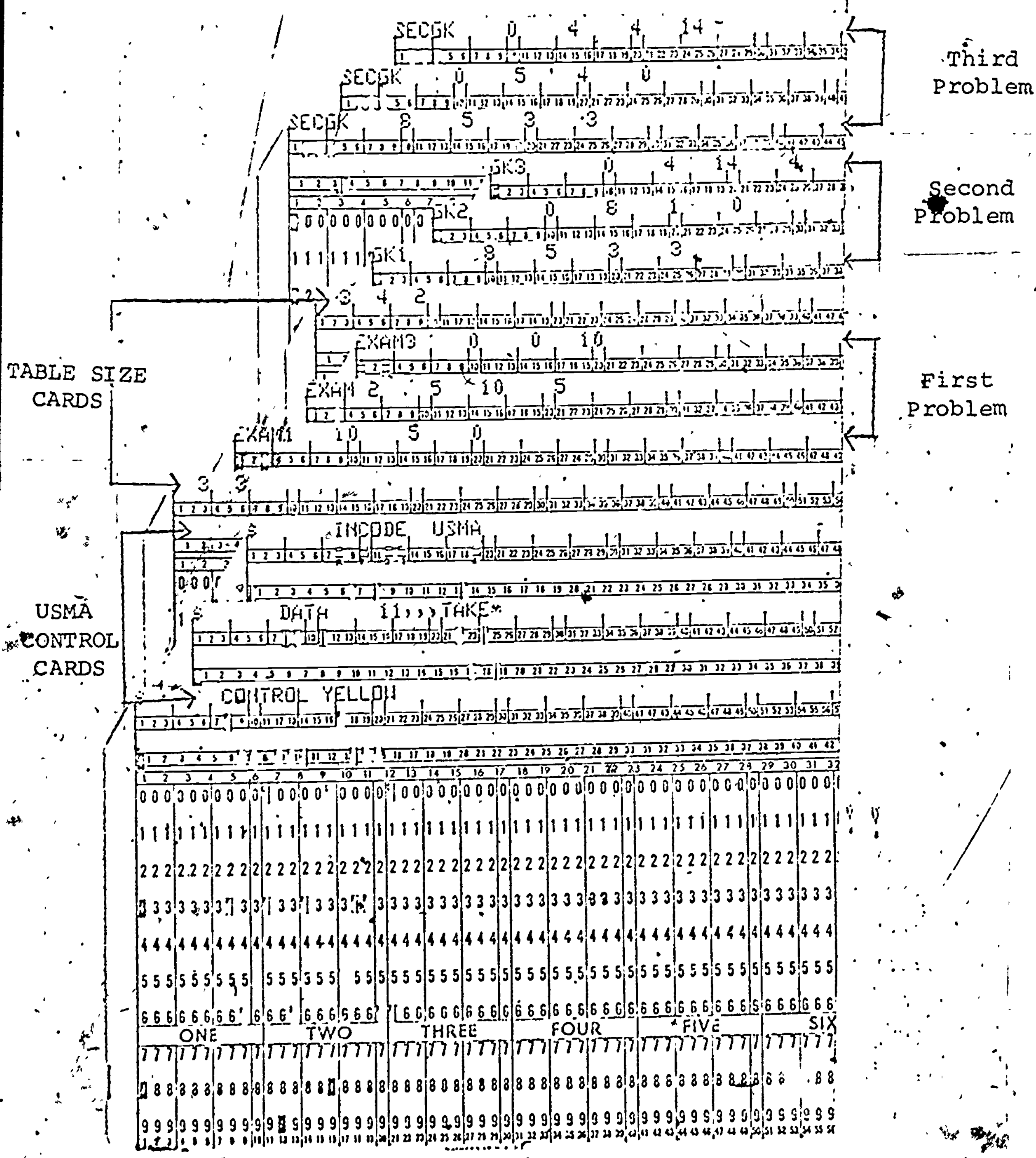
(2)

```
      PDD=0.
      PSD=0.
      DO 32 I=1,NR
       DO 32 J=1,NC
C   PS IS THE NUMBER OF PAIRS OF PERSON S WHO ARE RANKED SIMILARLY ON BC
C  PD IS THE NUMBER OF DISSIMILAR PAIRS ON BOTH X AND Y VARIATES
      PS=PS+X(I,J)*S(I,J)
      PD=PD+X(I,J)*D(I,J)
      PSS=PSS+ S(I,J)*X(I,J)*S(I,J)
      PDD=PDD+ D(I,J)*X(I,J)*D(I,J)
      PSD=PSD+S(I,J)*X(I,J)*D(I,J)
   32 CONTINUE
      GAMMA=(PS-PD)/(PD +PS)
C     LARGE SAMPLE ASYMPTOTIC STANDARD ERROR
      ZEED=PS *PDD*PS -2.0*PS*PD*PSD  +PD*PSS*PD
      ZEET=SQRT(ZEED)
      ZEE=(PD+PS)**2/(4.0*ZEET)
      IF(    (TX/RC).GE.4.0) GO TO 34
       PRINT 35
   35 FORMAT( " SAMPLE SIZE PER CELL IS NOT LARGE ENOUGH FOR "
     X    " MOST POWERFUL ASYMPTOTIC TEST." )
      IF(   (TX/RC).LT.2.0) GO TO 37
C   A MORE CONSERVATIVE STANDARD ERROR FOR SMALLER SAMPLES
      ZEED=(PS+PD)/(2.0*TX*(1.0-GAMMA**2 ))
      ZEE=      SQRT(ZEED)
   34 CONTINUE
      ERR=1.0/ZEE
      PRINT 105,GAMMA,ERR
  105 FORMAT( " SAMPLE GAMMA = ",F8.4," STANDARD ERROR ",F8.4)
      UL     =GAMMA+1.96*ERR
      WL     =GAMMA-1.96*ERR
      PRINT 106,UL,WL
  106 FORMAT( " PROBABLY(.95) POPULATION GAMMA IS BETWEEN "
     XF8.4, " AND ",F8.4)
      GO TO 401
   37 PRINT 38
   38 FORMAT( " THE SAMPLE SIZE PER CELL IS TOO SMALL")
      PRINT 402,GAMMA
  402 FORMAT(1H ,"GAMMA IS ", F10.5)
  401 CONTINUE
      GO TO 12
   99 STOP
      END

RE , NO   DIAGNOSTICS IN ABOVE COMPILATION
DS WERE USED FOR THIS COMPILATION
```

(3)

PHYSICAL PICTURE OF CONTROL CARDS
AND DATA CARDS FOR
SAMPLE PROBLEMS



c. Example of Program Output

```
EXAM1      10,0      5,0      0,
EXAM2       5,0     10,0      5,0
EXAM3       0,       5,0     10,0
SAMPLE GAMMA =    0,8333 STANDARD ERROR    0,0645
PROBABLY(.95) POPULATION GAMMA IS BETWEEN    0.9599 AND    0.7068
GK1         8,0      5,0      3,0      3,0
GK2         0,       8,0      1,0      0,
GK3         0,       4,0     14,0      4,0
SAMPLE GAMMA =    0,6122 STANDARD ERROR    0,1510
PROBABLY(.95) POPULATION GAMMA IS BETWEEN    0,9082 AND    0,3161
SECGK       8,0      5,0      3,0      3,0
SECGK       0,       5,0      4,0      0,
SECGK       0,       4,0      4,0     14,0
SAMPLE GAMMA =    0,7062 STANDARD ERROR    0,1158
PROBABLY(.95) POPULATION GAMMA IS BETWEEN    0.9333 AND    0.4792
```

*NOTE: This table, and the GAMMA associated, is the table used by Goodman and Kruskal (1963).

5. References:

Goodman, L. & Kruskal, Wm. "Measures of association for cross classification: III approximate sampling theory" Journal of the American Statistical Association 1963, 58, 310-364

Kendall, M. G. Rank Correlation Methods, Fourth Addition; London: Griffin, 1970.

Rosenthal, I. "Distribution of the sample version of the measure of association, GAMMA" Journal of the American Statistical Association, 1966, 61, 440-453.

ROBERT F. PRIEST, Ph.D.
Research Branch
Office of the Director of
Institutional Research

Robert F Priest